
Feature Extraction in Satellite Imagery Using Support Vector Machines

Kevin Culberg¹ Kevin Fuhs¹

Abstract

Satellite imagery is collected at an every increasing pace, but analysis of this information can be very time consuming. Analysts are typically required to review and label individual images by hand in order to identify key features. We set out to automate some of this analysis by classifying pixels within satellite images as one of eight different classes including man-made and natural features such as buildings, crops, and trees. We trained support vector machine (SVM) and logistic regression models to predict features classes from multi spectral satellite imagery. We found that logistic regression performed better on the Jaccard score than SVM with both models performing better on crops, waterways, and buildings then the other classes.

1. Introduction

Thanks to improvements in technology, the collection and availability of satellite imagery has exploded in recent years. Fortunately, this grants us unprecedented levels of data from areas with limited ground reporting. However, the challenge exists to translate this imagery into useful information (for the purposes of aid distribution, poverty identification, etc.), which is where the insights of machine learning are essential. To aid in this, one needs to recognize different man-made and natural features of an area such as buildings, waterways, roads, crops, etc. Our goal is to generate feature masks given satellite imagery that map each pixel to one of eight classes. The resulting mask can then be used for further analysis in a wide variety of applications, such predicting demographic data for use in projects such as aid distribution and finding optimal location of new markets.

¹Stanford University, Stanford, California, USA. Correspondence to: Kevin Culberg <kculberg@stanford.edu>, Kevin Fuhs <kfuhs@stanford.edu>.

2. Related Work

The literature review we conducted focused on the different ways to use support vector machines in the classification of imagery. The original motivation to use support vector machines and logistic regression models came from sources such as Jorge Ingladas article Automatic recognition of man-made objects in high resolution optical remote sensing images by SVM classification of geometric image features which showed success in making classifications with respect to satellite imagery with RGB and infrared pixels (Inglada, 2007). Furthermore, a key component of this was with respect to pixel classification and the expansion of the feature space, as done through the Gaussian and Laplacian filters. The motivation for this approach was based on research which found that multi-feature models produced by this technique have a positive impact on the overall model performance, including filters such as Gaussian and Laplacian. Additional research comparing different techniques to use binary SVM models for multi class classification problems to other techniques provided additional details (Melgani & Bruzzone, 2004; Foody & Mathur, 2004). We ended up employing the one vs all method to train only as many models as we had classes and using the prediction with the highest confidence. Other sources of research where we derived our motivation to use feature masks came from Huang and Zhang, 2012, who used the data to predict feature masks via image segmentation, which gave useful information about evaluation functions such as the Jaccard score which were useful in comparing the true labeled data to the mask generated by the model (Huang & Zhang, 2013). We further reviewed the various techniques for SVM in the meta-analysis conducted by Mountrakis (Mountrakis et al., 2011). Some of the ideas we tested included various kernel shapes and forms of parameter regularization - however, many of the techniques were not applicable to our model such as fuzzy clustering, since each point can only belong to one feature class in our labeled data. However, some techniques were used such as weighting of low-representation classes in order to improve training.

3. Dataset and Features

As our model input we used multi spectral satellite imagery from Kaggle dataset DSTL Satellite Imagery Feature Detection (Kaggle). The dataset consisted of labeled satellite images which averaged 800 by 800 pixels in size. Of these images, 16 contained a diversity of feature classes that made them useful for training our models. We sampled 1600, 64x64 pixel sub images for training and validation and 400 sub images for testing. The satellite imagery consists of sixteen bands in the visible and other wavelengths, of which we are using eight. The eight bands used by our models include red, red edge, coastal, blue, green, yellow, and two bands of near infrared. Image resolution varies from .31 to 7.5 meters per pixel. The feature space was expanded by applying Gaussian and Laplacian filters with a kernel size of 15 by 15 pixels to each band of the training data in order to reduce noise and find edges in images. These filters had the added benefit of incorporating data from nearby pixels in the available features. Each feature in both the training and test sets were normalized to have zero mean and a standard deviation of one using the mean and standard deviation from the training set. Figure 1 contains an example of the red, green, and blue bands of one input satellite image from this dataset.



Figure 1. Example satellite image used for model input. Only the red, green, and blue bands are shown.

The Kaggle dataset also included very detailed true class labels for each satellite image (Kaggle). Classes included buildings, other man-made structures, roads, dirt tracks, trees, crops, waterways, and standing water, shown in Figure 2.

These labels were included as shapefiles that we scaled and converted to a mask with the same height and width of the input image where each pixel value corresponds to one of the eight classes and a value of zero corresponds to no class present. An example of a class mask is shown in Figure 3.

1 – Building		5 – Trees	
2 – Man-Made Structures		6 – Crops	
3 – Roads		7 – Waterway	
4 – Dirt Tracks		8 – Standing Water	

Figure 2. Legend denoting each feature class and their corresponding color.

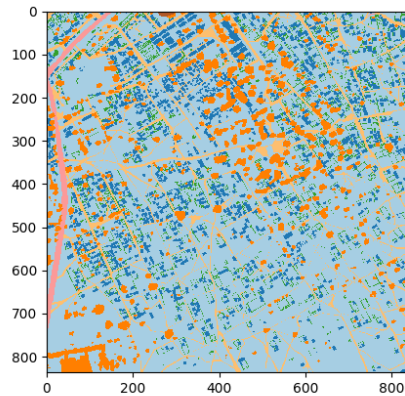


Figure 3. True class labels for the example input image.

4. Methods

We approached the model as a classification problem, attempting to predict the class for each pixel in an image. To increase the number of features and incorporate neighboring pixel information we applied Gaussian and Laplacian filters. The Gaussian filter was used to remove any noise from the satellite image while the Laplacian filter was used to highlight edges. After applying these filters we ended up with 24 features for each pixel from the original eight image bands. These features were used as input to train SVM and logistic regression models to predict a class for each pixel and generate a class mask with the same height and width as the input image.

To measure the success of each model we calculated both the accuracy and Jaccard score. For accuracy we calculated the recall which is found by dividing the true positives by the true positives and false negatives. We also calculated

the Jaccard score in order to take into account a models tendency to over classify for some classes.

$$J(A, B) = \frac{|A \cap B|}{|A \cup B|} = \frac{|A \cap B|}{|A| + |B| - |A \cap B|}$$

The Jaccard score (seen above) is the area of intersection over the area of the union for two sets. We calculate the Jaccard score for each class in order to take into account over prediction of a class when measuring performance.

In order to use our SVM and logistic regression models for the multi-class classification problem we trained nine versions (eight for each class and one for no class) of each model using the one vs all approach. In this approach we train each version of the model to be a binary classifier as either a member of its class or not. To obtain a final predicted class, the confidence from each of these models is compared to determine the final predicted class. Some of the classes had lower representation in the training data than others such as standing water and man-made structures not buildings. In order to improve model accuracy on these rare classes we weighted each class inversely to how common they were in the training data.

The SVM model aims to create a classification hyperplane and maximize the distance between labeled data using the squared hinge loss. We used an l2 regularization parameter and a linear kernel.

$$\begin{aligned} \min_{w, b, \zeta} \frac{1}{2} w^T w + C \sum_{i=1}^n \zeta_i \\ \text{subject to } y_i(w^T \phi(x_i) + b) \geq 1 - \zeta_i, \\ \zeta_i \geq 0, i = 1, \dots, n \end{aligned}$$

The goal was to create a model which found across multiple classes the strongest separating hyperplanes to predict various feature classes. Note that because this is multi-class classification we choose the class for each point which produces the greatest margin as calculated above. Logistic regression is based on the function $h_{\theta}(x) = g(\theta^T x) = 1/(1 + \exp(-\theta^T x))$ with the added l2 regularization parameter would give us the problem minimize $\sum_{x, y} \log(1 + \exp(w^T xy)) + w^T w$ The logistic regression function penalizes additional weights for each feature band and otherwise works to minimize the difference between the model-produced mask versus the actual feature values. Cross-entropy loss is used for logistic regression. All models were implemented using the Scikit-learn Python library (Pedregosa et al., 2011).

5. Experiments/Results/Discussion

For our experiments we trained an SVM and logistic regression model on 1600 64x64 pixel satellite images and used 400 images as our test set. From our experiments we found multiple interesting results. We found that logistic regression performed slightly better than SVM on the Jaccard score for all classes, Figure 4. Results were mixed for model accuracy between the two models, Figure 5. In addition, we identified significant variation in both Jaccard score and accuracy across classes for both models. Classes where our models underperformed tended to be those that had lower representation in our training data such as roads, standing water, and other man-made structures not buildings. These classes were also similar to other classes and relied on contextual clues such as shape to be able to properly infer the correct class. Despite our attempts to weight the training on these features our models still underperformed. Also, a significant discrepancy between model accuracy and Jaccard scores existed in certain feature classes, most notably standing water and roads. Practically, this means that while the model performed reasonably well at avoiding false negatives, false positives were a significant problem and the model over-classified on those feature classes.

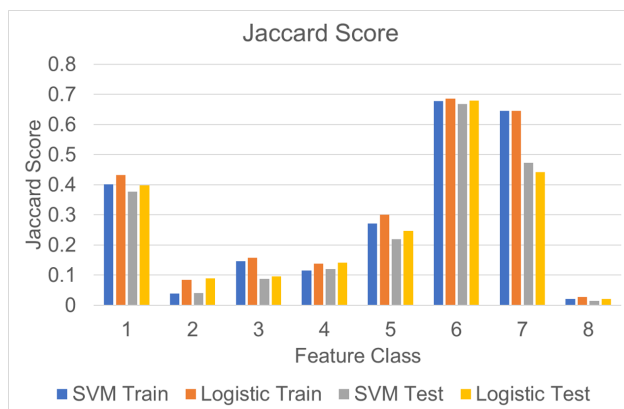


Figure 4. Train and test set Jaccard score for the SVM and logistic regression model across all classes. Classes include: 1 - buildings, 2 - man-made structures, 3 - roads, 4 - dirt tracks, 5 - trees, 6 - crops, 7 - waterways, and 8 - standing water.

Both the SVM and logistic regression models over classified pixels as trees and crops producing some bleed over onto nearby pixels on their output features masks, Figures 6 and 7. It is likely that our model was not able to learn the difference between an open grassy area and a crop given the limited training set. Additionally, it appears that our models do far better at identifying surface material and are less accurate at using the shapes in the image to predict the class. For example, on the left side of the satellite image

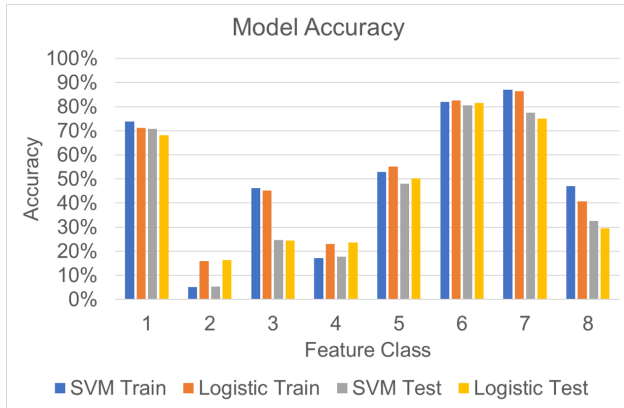


Figure 5. Train and test set accuracy for the SVM and logistic regression model across all classes. Classes include: 1 - buildings, 2 - man-made structures, 3 - roads, 4 - dirt tracks, 5 - trees, 6 - crops, 7 - waterways, and 8 - standing water.

used as input, Figure 1, there is a road that runs north-south. Both SVM and logistic regression model classify this area as a dirt track because the road surface is dirt and the model is unable to recognize that it is the width of the road that makes it a road. Another example is how some of the buildings in the north of the image are misclassified as road because of the roof material. If our models were able to recognize the shape of these structures then it has a higher accuracy in classifying these types of man-made objects.

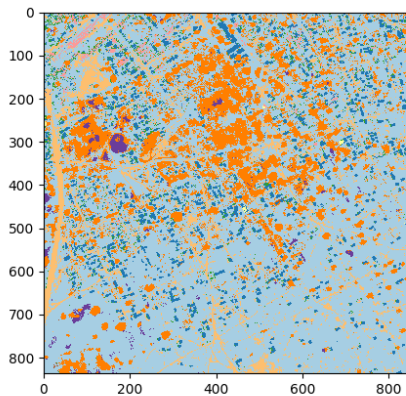


Figure 6. Predicted class labels from the logistic regression model.

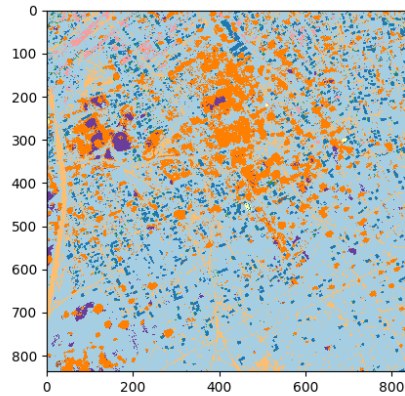


Figure 7. Predicted class labels from the SVM model.

6. Conclusion/Future Work

Support vector machines and logistic regression can be effective ways to classify man-made and natural features in satellite imagery. Through the use of filters to add features, the weighting of low representation classes, data normalization, and other factors, the models can predict many of the classes with a high degree of accuracy and low rate of false positives. We found that multi-class logistic regression performed better than a support vector machine using a linear kernel on all feature classes using Jaccard score as the metric. There was significant variability across different feature classes with features such as roads and tracks having significantly lower model accuracy than waterways and crops. We attribute this discrepancy with the low frequency of these classes in the training data. For future investigation, we would like to improve the classification rate for roads and dirt tracks by adding a line detection feature to the data. Additional improvements to classification accuracy could also be obtained by training separate models for each type of class with features that are tailored to identification of that specific class. Training these models would likely improve accuracy of classification on the classes that rely more on shape and structure of nearby pixels such as buildings and other man-made features.

7. Contributions

Kevin Culberg designed, implemented and tested all experiments as well as writing the functions for data manipulation, visualizations, the support vector machine model and other helper functions. Kevin Fuhs assisted by providing the initial implementation for the logistic regression model and feedback on experiments. All team members contributed to the research of different techniques to the

project problem as well as drafting all written project deliverables.

References

- Foody, G. M. and Mathur, A. A relative evaluation of multiclass image classification by support vector machines. *IEEE Transactions on Geoscience and Remote Sensing*, 42(6):1335–1343, June 2004. ISSN 0196-2892. doi: 10.1109/TGRS.2004.827257.
- Huang, X. and Zhang, L. An svm ensemble approach combining spectral, structural, and semantic features for the classification of high-resolution remotely sensed imagery. *IEEE Transactions on Geoscience and Remote Sensing*, 51(1):257–272, Jan 2013. ISSN 0196-2892. doi: 10.1109/TGRS.2012.2202912.
- Inglada, Jordi. Automatic recognition of man-made objects in high resolution optical remote sensing images by svm classification of geometric image features. *ISPRS Journal of Photogrammetry and Remote Sensing*, 62(3):236 – 248, 2007. ISSN 0924-2716. doi: <https://doi.org/10.1016/j.isprsjprs.2007.05.011>. URL <http://www.sciencedirect.com/science/article/pii/S092427160700055X>.
- Kaggle. Dstl satellite imagery feature detection. URL <https://www.kaggle.com/c/dstl-satellite-imagery-feature-detection>.
- Melgani, F. and Bruzzone, L. Classification of hyperspectral remote sensing images with support vector machines. *IEEE Transactions on Geoscience and Remote Sensing*, 42(8):1778–1790, Aug 2004. ISSN 0196-2892. doi: 10.1109/TGRS.2004.831865.
- Mountrakis, Giorgos, Im, Jungho, and Ogole, Caesar. Support vector machines in remote sensing: A review. *ISPRS Journal of Photogrammetry and Remote Sensing*, 66(3):247 – 259, 2011. ISSN 0924-2716. doi: <https://doi.org/10.1016/j.isprsjprs.2010.11.001>. URL <http://www.sciencedirect.com/science/article/pii/S0924271610001140>.
- Pedregosa, F., Varoquaux, G., Gramfort, A., Michel, V., Thirion, B., Grisel, O., Blondel, M., Prettenhofer, P., Weiss, R., Dubourg, V., Vanderplas, J., Passos, A., Cournapeau, D., Brucher, M., Perrot, M., and Duchesnay, E. Scikit-learn: Machine learning in Python. *Journal of Machine Learning Research*, 12:2825–2830, 2011.